



FUNDAÇÃO  
GETULIO VARGAS

**EPGE**

Escola de Pós-Graduação  
em Economia

## Ensaio Econômico

Escola de

Pós Graduação

em Economia

da Fundação

Getúlio Vargas

Nº 552

ISSN 0104-8910

***Purchasing power parity and the unit root tests: A robust  
analysis***

***Zhijie Xiao  
Luiz Renato Lima***

*Julho de 2004*

# Purchasing Power Parity and the Unit Root Tests: A Robust Analysis

Zhijie Xiao\*

University of Illinois at Urbana-Champaign

Luiz Renato Lima†

Getulio Vargas Foundation

July 2, 2004

## Abstract

Empirical evidence suggests that real exchange rate is characterized by the presence of near-unity root and additive outliers. Recent studies have found evidence in favor of PPP reversion by using the quasi-differencing (Elliott et al., 1996) unit root test (ERS), which is more efficient against local alternatives but is still based on least squares estimation. Unit root tests based on least squares method usually tend to bias inference towards stationarity when additive outliers are present. In this paper, we incorporate quasi-differencing into M-estimation to construct a unit root test that is robust not only against near-unity root but also against nonGaussian behavior provoked by additive outliers. We re-visit the PPP hypothesis and found less evidence in favor of PPP reversion when non-Gaussian behavior in real exchange rates is taken into account.

## 1 Introduction

Testing the PPP hypothesis using the augmented Dickey-Fuller (ADF) unit root tests has produced little evidence of PPP reversion. On the other hand, new approaches based on panel data and multivariate tests have yielded strong support for the PPP hypothesis. This new approach suggests that unit root tests based on both cross-sectional and time-series data are more powerful than those based on time series observations. Recent criticism about this view has already emerged. First, Maddala et al (2000) analyzes many aspects of panel data unit root tests and conclude that it cannot rescue the PPP theory. Second,

---

\*Department of economics, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA. Email: zxiao@uiuc.edu.

†Graduate School of Economics, Getulio Vargas Foundation. Praia de Botafogo 190. Rio de Janeiro-RJ, Brazil. Email: luizr@fgv.br.

Taylor (2002) argues that multivariate unit root test may suffer from "missing middle", meaning that the such tests fail to recognize many intermediate cases in which some series are stationary and some are not. He also argue that even those multivariate tests that do not suffer from missing middle, they may still suffer from the weak power problem identified by M.P. Taylor and Sarno (1998).

If panel data and multivariate unit root tests cannot give the right answer to the problem of low power of univariate unit root tests, then how can we test efficiently the purchasing power parity hypothesis? A possible solution is to try to gain higher power using Quasi-Differencing based unit root tests such as in Elliot et al. (1996). The unit root tests suggested in Elliot et al. (which we will shall refer as the ERS test) has high power against root near unity, which is an important feature exhibited by real exchange rates. Taylor (2002) used a secular sample of 23 countries with observations ranging from 1892 to 1996 and found strong evidence in favor of PPP by using the ERS test. Cheung and Lai (1998) considered post-Bretton Woods data and found more support for the PPP when the ERS is used in place of ADF tests.

The traditional ADF test and the efficient ERS test are based on least squares estimation. In the absence of Gaussianity, the least square based procedures are generally less efficient than more robust methods. Many applications in nonstationary economic time series involve data that are affected by infrequent but important events such as oil shocks, wars, natural disasters, and changes in policy regimes, indicating the presence of nonGaussian behavior in macroeconomic time series (see Balke and Fomby, 1994). It is well-documented that financial time series such as interest rate and exchange rate have heavy-tailed distributions. In such cases, it is important to consider estimation and inference procedures that are robust to departures from Gaussianity and can be applied to nonstationary time series. For the particular case of real exchange rates, empirical evidence suggests that they are plagued with additive outliers (AO) which moves unit root inference based on least squares towards stationarity (Hoek et al., 1995).

In this paper, we will explore a new approach that is less sensitive to AO, has more power than OLS based unit root test under nonGaussian behavior and still retains the good power of the ERS test for roots near unity. In particular, we investigate the PPP reversion by taking into account two common characteristic in real exchange rate: near-unity root and additive outliers. We incorporate quasi-differencing into M-estimation to achieve efficiency gain in the presence of near-unit root and non-Gaussianity. To select appropriate criterion function in M-estimation, we consider unit root tests coupled with partially adaptive estimation. In this sense, the present paper try to provide an intermediate class of unit root testing procedures that are more efficient and robust than the traditional OLS based-test in the presence of heavy-tailed distributions and, on the other hand, simpler than unit root tests based on fully adaptive estimation using nonparametric methods. Although fully adaptive estimator has the theoretically attractive property of asymptotic efficiency, as suggested by Bickel (1982, p.664), partially adaptive estimation is a more practical goal because it avoids the difficulty of nonparametric estimation of score functions (also see similar ar-

guments in Potscher and Prucha (1986), Hogg and Lenth (1984), McDonald and Newey (1988), and Phillips (1994)). We consider a partially adaptive estimator based on the family of student- $t$  distributions (Postcher and Prucha 1986). The choice of this family is based on two facts: (i) the family of student- $t$  represents an important dimension of the space of distributions, including the normal distribution as a limiting case and the Cauchy distribution as a special case. Its adaptation parameter will depend on the scale and thickness parameters, which can be easily estimated from the data using the approach proposed by Potscher and Prucha (1986) and; (ii) unlike the OLS estimator, the M-estimator based on student- $t$  distribution has a bounded influence function, meaning that it is less sensitive to the presence of additive outliers.

Monte-Carlo experiments are conducted to investigate the finite-sample performance of the partially adaptive M-detrended test. Comparing to the conventional ADF and ERS tests, the Monte Carlo results indicate that there is little loss in using the proposed unit root test when the innovations are Gaussian, and the power gains from using our M-detrended test is substantial when there are outliers or non-Gaussian innovations.

We next consider two samples of real exchange rate cited in the literature as evidence of PPP reversion obtained through the use of the ERS test. The first one was used in Taylor (2002) and the second was considered by Cheung and Lai (1998). Inference based on our partially adaptive M-detrended test yields less evidence of PPP reversion than that obtained using the non robust ERS test. Our results suggest that taking into account outliers in real exchange rates can lead to rejection of the, otherwise nonrejected, PPP hypothesis.

This paper is organized as follows: Section 2 reviews the literature on PPP hypothesis. Section 3 presents the econometric model. In particular we study a joint estimation of the deterministic trend and the autoregressive coefficient in a model with general disturbance distributions. Asymptotic analysis of the M-estimator is provided and a M-detrended (demeaned) unit root test is proposed based on these asymptotics and joint estimation. Some Monte Carlo results is presented in section 4, indicating a power gain of the proposed unit root test in the presence of heavy-tailed distributions. In Section 5, we apply these techniques to testing the PPP hypothesis. Section 6 concludes.

A word on notation: the symbol  $\Rightarrow$  signifies weak convergence,  $\equiv$  signifies equality in distribution, and  $:=$  signifies definitional equality.  $L$  denotes lag operator. The expression  $\Delta = 1 - L$  is the difference operator, and  $\Delta_c$  signifies quasi-difference, which is defined by  $\Delta_c = 1 - (1 + c/n)L$ . The term  $I(k)$  denotes integration of order  $k$ . All limits are taken as  $T \rightarrow \infty$ , unless otherwise specified.

## 2 Testing the PPP Hypothesis: A Review

The PPP hypothesis has been the object of intense study in past works. It suggests that national price levels of two countries are equalized when expressed in the common currency unit. Even though the relation is not likely to hold in the short run due to price stickiness, the relation may hold in the long run. In

past studies, a popular way of testing the PPP hypothesis in the long run has been examining the statistical properties of the real exchange rate. The real exchange rate, by definition, is the relative national price levels of two countries expressed in common currency units. If the real exchange rate is stationary, deviations from the PPP are transitory and a mean reversion occurs. In such a case, the relative national price levels is equalized in the long run and the PPP hypothesis holds. However, if the real exchange rate has a unit root, deviations from the PPP are permanent and a mean (trend) reversion does not occur. In such a case, the PPP does not hold.

Many past studies have examined statistical properties of the real exchange rate to draw the implications on the empirical relevance of the purchasing power parity (PPP) in the long run. Earlier studies found less favorable evidence of PPP in industrial countries for the post-Bretton Woods period, by reporting that the null hypothesis of unit root against stationarity of the real exchange rates was not rejected. These tests were mostly implemented through the augmented Dickey-Fuller (ADF) and/or Phillips-Perron (PP) statistics. As pointed out by Frankel (1986), Froot and Rogoff (1995), and Rogoff (1996), a popular explanation for the failure to reject the unit-root null has been the low power of the unit-root tests, which is aggravated by the high persistency (a root near unity) of the real exchange rate and the short time-span of the post-Bretton Woods sample. Many past researchers have tried to resolve this low power problem by suggesting two possible approaches: 1) extending data points, by comprising both pre and post-Bretton Woods observations; and 2) employing panel data and multivariate unit root tests to exploit both cross-sectional and time-series variation.

The problem with the first option is that the sample covers periods of floating and fixed exchange rate regimes. The panel data and multivariate unit root tests, on the other hand, allow to extend the sample by pooling across different exchange rates during the post-Bretton Woods period. This has made these tests popular in testing the PPP hypothesis. We next provide a brief review of the use of panel data and multivariate unit root tests in PPP hypothesis testing.<sup>1</sup>

## 2.1 Panel Data Unit Root Tests With Homogeneous Rate of Mean Reversion

In order to test the PPP hypothesis by exploring cross-section and time series variation, Abuaf and Jorion (1990) considered the following model

$$q_{it} = \alpha_o + \alpha_1 q_{i,t-1} + u_{it} \quad (1)$$

where  $q_{it}$  is the (logged) real exchange rate of country  $i$  at time  $t$ . Under the null hypothesis  $H_o : \alpha_o = 0, \alpha_1 = 1$  the PPP hypothesis does not hold. They estimated the system by using pooling GLS and reject the null for the period

---

<sup>1</sup>Our review on panel data unit root testing is strongly based on Maddala (2000).

1900-72, but they found weak evidence in favor of PPP for the period post-Bretton Wood. The main problem associated to this test procedure is that it ignores possible serial correlation in  $u_{it}$  and also assume the same autoregressive parameters across  $i$ .

The first unit root test to use a panel model with specific effects and specific autoregressive coefficients was suggested by Levin and Lin (1992, 1993) and Quah (1994). The Levin and Lin (LL hereafter) test is based on the following model

$$\Delta q_{it} = \alpha_i q_{i,t-1} + c_i + \varepsilon_{it} \quad (2)$$

where  $c_i$  is a country specific effect and  $\varepsilon_i$  allowed to display  $I(0)$  serial autocorrelation. In order to construct the test statistic, LL estimate the ADF regression below

$$\Delta \bar{q}_{it} = \alpha_i \bar{q}_{i,t-1} + \sum_{j=1}^k \theta_{ij} \Delta \bar{q}_{i,t-j} + c_i + e_{it} \quad (3)$$

where  $\bar{q}_{it}$  is the (logged) real exchange rate in country  $i$  at time  $t$  after subtraction of the cross-section average. Then, they perform two auxiliary regressions of  $\Delta \bar{q}_{it}$  and  $\bar{q}_{i,t-1}$  on the remaining part of Eq. 3 and collect their respective residuals, say,  $\hat{\varepsilon}_{it}$  and  $\hat{v}_{i,t-1}$ . Since there may be heteroskedasticity in  $e_{it}$ , they suggest to use the normalized residuals, say,  $\tilde{\varepsilon}_{it}$  and  $\tilde{v}_{i,t-1}$ . The test statistic is obtained from the following panel residual based regression:

$$\tilde{\varepsilon}_{it} = \alpha \tilde{v}_{i,t-1} + e_{it} \quad (4)$$

They then show that the  $t_\alpha$  statistic has a standard normal distribution. The LL test has been largely applied to test the PPP hypothesis ( see, for example, MacDonald 1996, Frankel and Rose 1996, and Oh 1996). It is mostly concluded that panel analysis provides strong support in favor of PPP reversion because panel unit root tests have much more power than univariate unit root tests since the former is based on a pooled sample and, therefore, avoid the characteristic low power of univariate tests caused by short samples of the post-Bretton Woods era. However, as argued by Maddala et al (2000), this argument is nonsense because that the set of hypotheses of univariate and panel unit root tests are totally different from each other and, therefore, we cannot compare their powers. In other words, the unit root test for a single time series is:  $H_o : \rho_1 = 1$  vs  $H_1 : \rho_1 < 1$  in the single regression  $y_t = \rho_1 y_{t-1} + e_t$ . The panel data unit root test, on the other hand, tests the hypothesis

$$\begin{aligned} H_o & : \rho_1 = \rho_2 = \dots \rho_N = \rho = 1 \\ \text{vs } H_1 & : \rho_1 = \rho_2 = \dots \rho_N = \rho < 1 \end{aligned}$$

in the regression  $y_{it} = \rho_i y_{i,t-1} + e_{it}$ ,  $i = 1, 2, \dots, N$ . One can see that the two null hypothesis are different and we cannot compare the power of these tests.

Therefore, panel data unit root tests may be the wrong answer to the low power problem of univariate unit root tests.

Recent studies have also concluded that the LL test have high size distortions in the presence of serial correlation in the disturbance terms. In particular, Papell (1997) considers the following model

$$q_{jt} = \rho_j q_{j,t-1} + \sum_{i=1}^k \gamma_{ij} \Delta q_{j,t-i} + \alpha_j + e_{jt} \quad j = 1, \dots, N \quad (5)$$

He tabulates new critical values and show that the critical values in Levin and Lin may lead to inappropriate rejections of the unit root hypothesis. Using the finite critical values he computed, Papell applies the LL test to six samples and rejects the unit root null at 5% for only some of his samples but not all of them.

O'Connel (1998) studied the importance of cross-section correlation in the disturbances. He concludes that there are dramatic size distortions for the LL test in the presence of contemporaneous correlation. Overall, the works of Papell and O'Connel seems to suggest that the strong evidence of PPP reversion obtained via panel data unit root tests may have been spurious.

Another way to study the PPP hypothesis using panel models was suggested by the study of Frankel and Rose (1996). They estimate the following panel regression model using data on 150 countries over a period of 45 years.

$$\Delta q_{it} = \alpha + \delta q_{i,t-1} + \sum_i \beta_i D_i + \sum_t \phi_t D_t + \varepsilon_{it} \quad (6)$$

where  $q_{it}$  is the real exchange rate for country  $i$  at time  $t$ , and  $D_i$  and  $D_t$  are country specific and year dummies respectively. Based on regression 6, Frankel and Rose conclude that there is enormous evidence in favor of the PPP hypothesis and that this support relies strongly on cross-section data with the time series data masking favorable evidences of PPP. Maddala et al (2000), however, show that the arguments used by Frankel and Rose may be flawed in several respects. They first argue that although time series observations of real exchange rates are plagued with the problem of instability over time, there is also strong evidence of cross-sectional heterogeneity: the results change when countries are deleted from the sample or when different country groups are considered. In order to investigate the conclusion of Frankel and Rose which states that support for the PPP relies strongly on cross-sectional rather than time-series data, Liu and Maddala (1996) presented cross-sectional and time-series regressions. Looking at cross-sectional regressions, they find that except in one single case in the 42 years considered, the coefficient  $\delta$  was not significantly different from zero, implying that there is no evidence that cross-sectional data support the PPP hypothesis. In estimating the time-series regression, Liu and Maddala found support for the PPP using WW-II time-series data. The favorable evidence came from several African countries, Australia and New Zealand, and two Asian countries Thailand and South Korea. Based on these findings, Maddala et al (2000) conclude that: "It is not true, as concluded by Frankel and Rose,

that strong support for PPP can be masked in time series data but are relatively easy to find in cross-section data.” They then suggest to explore more efficient univariate unit root tests such as the one suggested by Elliot et al (1996).

## 2.2 Panel Data Unit Root Tests With Heterogeneous Rate of Mean Reversion.

Under the alternative hypothesis in the panel data unit root tests discussed above, each real exchange rate reverts to its respective mean at the same rate. This strong restriction has motivated many works to consider panel data unit root tests with heterogeneous rate of mean reversion. In this direction, Im et al (1996) propose the IPS in which the alternative hypothesis under study is  $H_1 : \rho_j < 1, j = 1, \dots, N$  and, therefore,  $H_1$  now allows different convergence rates for different countries. The procedure of the IPS test can be briefly summarized as follows. Let  $t_{\rho_j}$  be the ADF t-statistic for the  $j - th$  currency.

The IPS test statistic is defined as  $\bar{z} = \frac{\bar{t} - E(\bar{t})}{\sqrt{Var(\bar{t})}}$ , where  $\bar{t} = \frac{1}{N} \sum_{j=1}^N t_{\rho_j}$ , with mean  $E(\bar{t})$  and variance  $Var(\bar{t})$ . This test statistic is asymptotically normal. Monte Carlo experiments implemented by Im et al shows that IPS test has higher power than the LL test. Note that the IPS test just pools individual test statistics across countries. Maddala and Wu (1998) suggest a new approach (called Fisher test) which also allows different convergence rates under  $H_1$ . Their test pools p-values of each individual test across countries to calculate the Fisher test statistic  $\lambda = -2 \sum_{j=1}^N \log P_j$ , where  $P_j$  is the p-value ADF test for the  $j - th$  country. The Fisher statistics has  $\chi^2$  distribution with  $2N$  degrees of freedom. Monte Carlo experiments also confirm that the Fisher test has higher power than the LL test. Wu and Wu (1998) apply IPS and Fisher tests to test the PPP hypothesis. They find that both tests reject the null of a unit root strongly for all samples, which implies that the PPP holds according to the results from IPS and Fisher tests.

However, as argued by Taylor (2002), these tests suffer from a ”missing middle”: the null corresponds to  $N$   $I(1)$  time series, but the alternative implies that all  $N$  series are  $I(0)$ . Therefore, this structure is not appropriate to identify many intermediate cases in which some time series are stationary and some are not. Furthermore, these ADF tests based on panel data may reject the unit root null when only some of the series are stationary, but not all  $N$  series.

## 2.3 Multivariate Unit Root Tests.

In order to solve the problem of missing middle, M.S. Taylor and Sarno (1998) develop a multivariate unit root test in which the null hypothesis is violated only when all series are stationary.

Let  $q_t = (q_{1t}, \dots, q_{Nt})$  be the  $N \times 1$  vector of real exchange rates at time  $t$ . M.S. Taylor and Sarno consider the following model



$$\Delta q_t = \Gamma_1 \Delta q_{t-1} + \dots + \Gamma_{k-1} \Delta q_{t-k+1} + \Gamma_k \Delta q_{t-k} + \mu + \omega_t \quad (7)$$

which is just a error-correction representation for the dynamic of the real exchange rate. Under the PPP hypothesis, all  $N$  components of  $q_t$  are  $I(0)$ . The  $N \times N$  matrix  $\Gamma_k$  has rank equal to the number of cointegrating vectors so that the null hypothesis corresponds to one or more nonstationary series, that is  $H_o : \text{rank}(\Gamma_k) < N$ , and the alternative corresponds to all series being stationary, that is  $H_1 : \text{rank}(\Gamma_k) = N$ . The test statistic for this test is the Johansen likelihood ratio,  $\text{JLR} = -T \ln(1 - \lambda_N)$ , where  $\lambda_N$  is the smallest eigenvalue of  $\Gamma_k$  and  $T$  is the time series size. The JLR statistic has a limiting chi-square distribution with one degree of freedom.

The problem with this test is that it places further restrictions on the lag structure since it imposes  $k$  lags of the  $\Delta q_t$  in all the equations of the system<sup>2</sup>. Taylor (2002) applies this test and found weak support for the PPP hypothesis, which indicates that the JLR test may suffer of weak power. Based on this result, Taylor proposes exactly the same approach suggested by Maddala et al.: shifting attention to a univariate approach that uses the most possible efficient univariate test. He also recommend to use the ERS test which is briefly described next.

## 2.4 ERS Test

In so far, we have learned from the previous sections that panel data and multivariate unit root tests are not the right answer to the problem of low power of univariate unit root tests. In order to gain higher power, the recent literature has recommended to use the more efficient ERS test proposed by Elliot et al (1996). The ERS test is an univariate unit root test that has high power against roots near unity. Elliot et al. obtains the asymptotic power envelope for unit root tests analyzing the sequence of Neyman-Pearson tests of the null hypothesis  $H_o : \alpha = 1$  against the local alternative  $H_1 : \alpha = 1 + c/n$ . The version of the test that allows for a linear trend is based on the regression below:

$$(1 - L)y_t^\tau = \phi_o y_{t-1}^\tau + \sum_{j=1}^p \phi_j (1 - L)y_{t-j}^\tau + v_t \quad (8)$$

where  $v_t$  is an error term and  $y_t^\tau$  is the locally detrended data process under the local alternative of  $\alpha = 1 + c/n$ , where  $n$  is the sample size. This is illustrated by:

$$y_t^s = y_t - \bar{\gamma}' x_t \quad (9)$$

where  $x_t = (1, t)$  is a linear trend and  $\bar{\gamma}$  is the least square regression coefficient of  $\Delta_{\bar{c}} y_t$  on  $\Delta_{\bar{c}} x_t$ , where  $\Delta_{\bar{c}} y_t = [y_1, (1 - \alpha L)y_2, \dots, (1 - \alpha L)y_n]'$  and  $\Delta_{\bar{c}} x_t = [x_1, (1 - \alpha L)x_2, \dots, (1 - \alpha L)x_n]'$ . The ERS statistic is given by the t-ratio, when testing  $H_o : \phi_o = 0$  against  $H_1 : \phi_o < 0$  in regression 8. Elliot et al (1996) recommended the localizing parameter,  $c$ , to be set equal to  $-13.5$ . For the

<sup>2</sup>This restriction is not required in the Fisher and IPS tests.

specification without a linear trend but still with a constant, one should go over the same steps by replacing  $y_t^\tau$  with the locally demeaned series  $y_t^\mu$  and  $x_t = 1$ . For this case, one should now set  $\bar{c} = -7$ .

The importance of the ERS test is that it is robust against root near unity. Nevertheless, as we will show next, real exchange rates may be contaminated by additive outliers, which will move OLS based unit root test (such as the ERS test) towards stationarity.

## 2.5 Additive Outliers

In this section, we introduce the concept of additive outlier (AO) and show its effect on unit root inference. The discussion presented here is strongly based on the work of Hoek et al. (1995). Suppose that the stochastic process is generated according the following DGP

$$\begin{aligned} y_t &= (1 - z_t)x_t + z_t w_t \\ \phi(L)x_t &= \varepsilon_t \\ w_t &= x_t + \xi_t \end{aligned} \tag{10}$$

where  $\phi(L)$  is a polynomial of order  $p$  in the lag operator  $L^3$ ,  $z_t$  is a binary random variable, which is equal to 1 with probability  $\gamma$  and to 0 otherwise,  $x_t$  is the outlier-free process,  $w_t$  is the contaminating process, and  $\xi_t$  is the additive outlier. Therefore, model (10) tells us that the time series will be equal to  $x_t$  (the outlier-free process) with probability  $\gamma$  and equal to  $x_t + \xi_t$  otherwise. Figure 1 exhibits a realization with 200 observations of the process  $y_t$  for  $\phi(L) = 1 - L$ ,  $\xi_{70} = 15$  (meaning that the process  $y_t$  is contaminated at time  $t = 70$ ), and  $\{\varepsilon_t\}$  a set of iid normal innovations. It is evident that an additive outlier causes a single departure from the normal pattern of the time series: the series immediately returns to its normal trajectory after the shock provoked by the AO is over.

The effect of additive outliers on unit root inference can be shown through the influence function (IF). Roughly speaking, the IF measures the change in the value of an estimator when an outlier is added to the sample. As an example, consider the AR(1) process  $x_t = \phi x_{t-1} + \varepsilon_t$ , with  $t = 1, \dots, T$ , where  $\varepsilon_t$  is i.i.d. with finite variance. Let  $z_t = 0$  for all  $t \neq s$  and  $z_t = 1$  for some  $1 < s < T$ . Hoek et al. (1995) prove that the OLS estimator of  $\phi$  based on the observed series  $y_t$  will be equal to

$$\hat{\phi}_{ols} = \frac{\sum_{t=1}^T y_t y_{t-1}}{\sum_{t=1}^T y_{t-1}^2} = \frac{\xi(x_{s-1} + x_{s+1}) + \sum_{t=1}^T x_t x_{t-1}}{\xi^2 + 2\xi x_s + \sum_{t=1}^T x_{t-1}^2} = O_p(\xi^{-1}) \tag{11}$$

---

<sup>3</sup>Note that we do not restrict  $x_t$  to be stationary, that is, we can also assume that  $\phi(L) = 1 - L$

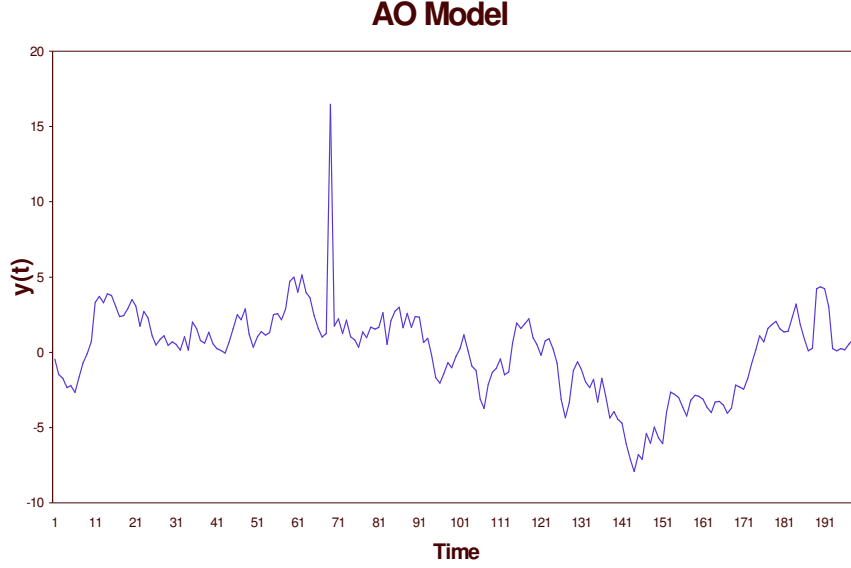


Figure 1: Effect of Isolated Additive Outlier

Note that  $\xi$  represents the additive outlier. Therefore, existence of  $\xi$  will make the OLS estimator,  $\hat{\phi}_{ols}$ , be biased towards zero. Hence, unit root tests based on OLS estimation, such as the ERS test, will likely overreject the null hypothesis of unit root when AO's are present.

Therefore, it is important to test for unit root taking into account not only near-unity alternatives (i.e., ERS test) but also the presence of additive outliers. The first step is to consider another estimator that possesses a bounded IF. A simple choice is the M estimator based on the class of Student-t distributions. This estimator is nothing else than a conditional maximum likelihood estimator within the class of iid Student-t distributed innovations (MLT). As an example, consider the AR(1) model  $y_t = \phi y_{t-1} + \varepsilon_t$ . The MLT estimator of  $\phi$  maximize the following log-likelihood function

$$L = \text{constant} + \frac{n}{2} \ln \Theta - \frac{\nu + 1}{2} \sum_{t=j+2}^n \ln \left\{ 1 + \frac{\Theta}{\nu} [\varepsilon_t]^2 \right\}$$

where the parameter  $\Theta$  measures the spread of the disturbance distribution and  $\nu$  is the degree of freedom that measures the tail thickness. Large values of  $\nu$  corresponds to thin tails in distribution. For given parameters  $\nu$  and  $\Theta$ , denoting  $\Theta/\nu$  as  $\theta$ , the MLT of  $\phi$  is the solution of the following optimization problem

$$\min_t \sum \varphi(\varepsilon).$$

where  $\varphi(\varepsilon) = \ln \left\{ 1 + \theta [\varepsilon_t]^2 \right\}$ . Hoek et al. (1995) noted that the criterion function,  $\psi(\varepsilon) = \varphi'(\varepsilon)$ , of the MLT estimator is  $O(\varepsilon_t^{-1})$  and therefore satisfies the necessary condition for the boundness of the influence function in a time series context. These results suggest that if one wants to conduct unit root inference robust against AO's, then it is relevant to consider unit root tests based on the MLT estimator. In the next section, we introduce a M-detrended unit root test, which will turn out to be robust against AO's as well as near unity alternatives.

### 3 The Econometric Model

We considered the following stochastic process:

$$y_t = \gamma' x_t + y_t^s, \quad t = 1, \dots, n. \quad (12)$$

$$y_t^s = \alpha y_{t-1}^s + u_t, \quad t = 1, \dots, n. \quad (13)$$

where the time series  $y_t$  is decomposed as the sum of a deterministic trend  $\gamma' x_t$  and a stochastic component  $y_t^s$ .  $y_t^s$  is the stochastic component which denotes the deviations of the series from the trend. These deviations are autocorrelated and the largest autoregressive root is  $\alpha$ .  $\{u_t\}$  is an  $I(0)$  process with mean zero.  $y_0$  is a constant (more generally, without affecting the asymptotic results, we can assume that it is a random variable of finite variance). This parameterization allows for the same deterministic component under both the null and the alternative hypothesis and is now widely used in economic time series analysis. Combining (12) and (13) gives the following nonlinear regression model.

$$y_t = \gamma' \Delta x_t + (1 - \alpha) \gamma' x_{t-1} + \alpha y_{t-1} + u_t. \quad (14)$$

In this paper, we study estimation and unit root testing for time series with general disturbance. We assume that  $\alpha$  may be close to unity and, therefore, reparameterize  $\alpha$  so that the parameter space is in a shrinking neighborhood of unity:

$$\alpha = 1 + c/n \quad (15)$$

Combining (12), (13), and (15), we have

$$\Delta y_t = \gamma' \Delta x_t - c \gamma' \left( \frac{x_{t-1}}{n} \right) + c \left( \frac{y_{t-1}}{n} \right) + u_t \quad (16)$$

In the simple case in which  $u_t$  is iid normally distributed, given observations on  $y_t$ , the maximum likelihood estimators of  $\gamma$  and  $c$  are the nonlinear least

squares estimators. In the absence of Gaussianity in  $u_t$ , it is possible to follow the idea of Huber (1964) for the location problem in order to obtain more robust estimators. In this direction, Relles (1968) and Huber (1973) introduced a class of so-called M estimators which generally have good properties over a wide range of distributions.

The M-estimator for  $(c, \gamma)$  in model (16) is defined as the solution of the following extreme problem:

$$\begin{bmatrix} \hat{c} \\ \hat{\gamma} \end{bmatrix} = \arg \max L(c, \gamma) \quad (17)$$

where

$$L(c, \gamma) = \sum_{t=2}^n \varphi(\Delta y_t - c(y_{t-1}/n) - \gamma' \Delta x_t + c\gamma'(x_{t-1}/n)).$$

for some criterion function  $\varphi$ . In the case that  $u_t$  are iid errors and  $\varphi$  is the true log density function of  $u$ ,  $L(c, \gamma)$  is the log likelihood function and the estimator  $(\hat{c}, \hat{\gamma})$  given by (17) is the maximum likelihood estimator.

The above analysis focuses on the major source of serial correlation in  $y_t^s$ , i.e.  $\alpha = 1 + c/n$ , and ignore the  $I(0)$  serial correlation in  $u_t$ . This generally brings no loss of asymptotic efficiency. However, it is possible to consider parameterizing the  $I(0)$  correlation in  $u_t$  using, say, ARMA(p,q) models.

Nonlinear estimation based on (17) is usually complicated. A relative simple and practically useful approach is to couple M-estimation with quasi-differencing (QD). Again, we are interested in testing the null hypothesis  $c = 0$  against the alternative  $c < 0$ . For some appropriate choice  $\bar{c}$ , we calculate the estimator for  $\gamma$  under the hypothesis  $c = \bar{c}$ ,

$$\bar{\gamma} = \arg \max \sum_{t=1}^n \varphi(\Delta \bar{c} y_t - \gamma' \Delta \bar{c} x_t). \quad (18)$$

The above quasi-difference based estimation removes the major (I(1)) correlation in time series  $y_t$ . Then, we construct the detrended data based on  $\bar{\gamma}$ ,

$$\bar{y}_t^s = y_t - \bar{\gamma}' x_t.$$

Re-estimating  $c$  based on the M-estimator of the autoregressive coefficients of  $\bar{y}_t^s$ , i.e.,

$$\hat{c} = \arg \max \sum_{t=1}^n \varphi(\Delta y_t - c(y_{t-1}/n) - \bar{\gamma}' \Delta x_t + c\bar{\gamma}'(x_{t-1}/n)), \quad (19)$$

The quasi-difference M-detrended unit root test (QDM) can then be constructed based on  $\bar{Z}_{QDM} = \hat{c}$ .

For  $\bar{\gamma}$ , it can be shown that (see appendix for asymptotic theory of M-estimator)

$$n^{-1/2} D_n(\bar{\gamma} - \gamma) \Rightarrow \left[ \int X_{\bar{c}}(r) X_{\bar{c}}(r)' dr \right]^{-1} \int X_{\bar{c}} d\bar{S}_{\bar{c}}(r) \quad (20)$$

where

$$\bar{S}_c(r) = S_0(r) - (\bar{c} - c) \int_0^r J_c, \quad S_0(r) = B_\varphi(r)/\delta, \text{ and } X_{\bar{c}} = g(r) - \bar{c}X(r).$$

Consequently, the partial sum process based on  $\bar{y}_t^s$  has the following asymptotic behavior:

$$\begin{aligned} \frac{1}{\sqrt{n}} \bar{y}_{[nr]}^s &= \frac{1}{\sqrt{n}} y_{[nr]}^s - \left[ (\bar{\gamma} - \gamma)' D_n \frac{1}{\sqrt{n}} \right] D_n^{-1} x_t \\ &\Rightarrow J_c(r) - \int d\bar{S}_c(r) X_{\bar{c}}' \left[ \int X_{\bar{c}}(r) X_{\bar{c}}'(r) dr \right]^{-1} X(r) \end{aligned} \quad (21)$$

We summarize the limiting null distribution and the power function of  $\bar{Z}_M = \hat{c}$  in the following proposition. For more discussion on related asymptotics, see Xiao (2001).

PROPOSITION:

(1) Under  $c = 0$ ,

$$\bar{Z}_{QDM} \Rightarrow \left[ \int \underline{B}_{\bar{c}}^2 \right]^{-1} \left[ \int \underline{B}_{\bar{c}} dS_0 - \int \underline{B}_{\bar{c}} g' \left[ \int X_{\bar{c}} X_{\bar{c}}' \right]^{-1} \int X_{\bar{c}} d\bar{S}_0 + \lambda \right]. \quad (22)$$

where

$$\underline{B}_{\bar{c}}(r) = B_1(r) - X(r)' \left[ \int X_{\bar{c}} X_{\bar{c}}' dr \right]^{-1} \int X_{\bar{c}} d\bar{S}_0(r), \text{ and } \bar{S}_0(r) = S_0(r) - \bar{c} \int_0^r B_1$$

(2) Under the alternative hypothesis,

$$\begin{aligned} \bar{Z}_{QDM} &\Rightarrow c + \left[ \int \bar{J}_{\bar{c}}^2 \right]^{-1} \left[ \int \bar{J}_{\bar{c}} dS_0 - \int \bar{J}_{\bar{c}} \underline{J}_c dr + \lambda \right], \\ \text{where } \underline{J}_c(r) &= g(r)' \left[ \int X_{\bar{c}} X_{\bar{c}}' dr \right]^{-1} \int X_{\bar{c}} d\bar{S}_c(r) - cJ_c(r), \\ \text{and } \underline{J}_{\bar{c}}(r) &= J_c(r) - X(r)' \left[ \int X_{\bar{c}} X_{\bar{c}}' dr \right]^{-1} \int X_{\bar{c}} d\bar{S}_c(r) \end{aligned}$$

To obtain asymptotically valid tests for a unit root, we need to know the limiting distributions. We may calculate the critical values by simulating Brownian Motions using Gaussian random variables. Notice that the Brownian motions  $B_\varphi(r)$  and  $B_1(r)$  are correlated, we may decompose  $B_\varphi(r)$  into a combination of orthogonal components  $B_1(r)$  and  $B_{\varphi,1}(r)$ . Notice that the parameters can be estimated consistently. The asymptotic null distribution is unaffected if the parameters are replaced by their consistent estimates. Thus, a robust estimate

of the null distribution can be obtained by simulating the distribution with the unknown parameters replaced by their consistent estimates. Alternatively, we may generate critical values under the null using re-sampling as follows: (1) First we detrend the data and obtain  $\bar{y}_t^s = y_t - \bar{\gamma}'x_t$ ; (2) Let  $\hat{u}_t = \Delta\bar{y}_t^s$  and fit  $\hat{u}_t$  by an  $AR(k)$  process with  $k \rightarrow \infty$  with  $n$ , but  $k = o(n^{1/3})$ , denote the estimated AR parameters as  $\{\hat{\theta}_j, j = 1, \dots, k\}$  and the residual of the fitted  $AR(k)$  process as  $\hat{\varepsilon}_t$ ; (3) Draw i.i.d.  $\{\varepsilon_t^*\}$  from  $\{\hat{\varepsilon}_t\}$ ; (4) Generate  $\{u_t^*\}$  from  $\{\varepsilon_t^*\}$  and  $\{\hat{\theta}_j, j = 1, \dots, k\}$ ; (5) Generate  $y_t^{s*}$  under the null of a unit root:  $y_t^{s*} = y_{t-1}^{s*} + u_t^*$ ; (6) Generate  $y_t^* = y_t^{s*} + \bar{\gamma}'x_t$ ; (7) Calculate the M-detrended unit root test  $Z_{QDM}^*$  based on  $y_t^*$ . Repeat the above process many times, each time obtain a calculated  $Z_{QDM}^*$  and the  $\tau$ -th quantile of these calculated  $Z_{QDM}^*$  may be used as the 100 $\tau$ % level critical value.

**Remark 1** (*Parameterizing  $I(0)$  Serial Correlation*) *The analysis of this paper focused first on the major source of serial correlation in  $y_t^s$ , i.e.  $\alpha = 1 + c/n$ , and ignored the  $I(0)$  serial correlation in  $u_t$ . This generally brings no loss of asymptotic efficiency. However, it is possible to consider parameterizing the  $I(0)$  correlation in  $u_t$  using, say,  $ARMA(p, q)$  models. For example, if we assume that the disturbance process  $u_t$  is  $AR(k)$  and  $y_t^s$  in (13) can be represented by*

$$y_t^s = \alpha y_{t-1}^s + \sum_{j=1}^k \psi_j \Delta y_{t-j}^s + \varepsilon_t, \quad (23)$$

where  $\varepsilon_t$  is an iid sequence, we may consider a M estimator of  $(\phi_o, \{\psi_j\}_{j=1}^k)$  that minimizes

$$\sum_{t=2}^n \varphi \left( \Delta y_t^s - \phi_o y_{t-1}^s - \sum_{j=1}^k \psi_j \Delta y_{t-j}^s \right).$$

The QDM unit root test can then be constructed based on  $\bar{Z}_{QDM} = \frac{n\hat{\phi}_o}{1 - \sum_{j=1}^k \hat{\psi}_j}$ . The asymptotic behavior of this statistic can be analyzed similarly.

### 3.1 Choosing the Criterion Function: A Partially Adaptive Estimator.

The M estimator for  $(c, \gamma)$  was motivated by the maximum likelihood estimator. In practice, even if the exact distribution of the innovations is unknown, if the data has similar tail behavior as the density function used in the estimation, then inference based on these method still have good sampling performance. Thus, in applications, we may consider adaptive or partially adaptive estimation methods so that the data density can be approximated.

Although from a theoretical point of view, fully adaptive estimation method based on nonparametric estimation (Hansen 1995; Seo 1996; Beelders 1998) is

ideal, partially adaptive estimation (Bickel 1982, pp664; Postcher and Prucha 1986) may be a practically useful method. In this section, we consider a partially adaptive M estimator for  $(c, \gamma)$ .

Taking into account of the well documented characteristic of heavy tails in economic and financial data, we consider a partially adaptive estimator based on the student-t distributions, although other classes of distributions may be analyzed similarly. The student-t distribution is an important class of distributions (see more discussion in, say, Hall and Joiner 1982) that contains the Cauchy distribution as a special case and the normal distribution as a limit case, and has wide applications in economic analysis. Partially adaptive estimator based on this class of distribution is reasonably robust.

The M-estimators considered in this section is also the MLE corresponding to the student-t distributed disturbances. In case of t-distributed innovations, the log-likelihood is given by

$$L(c, \gamma) = \text{constant} + \frac{n}{2} \ln \Theta - \frac{\nu + 1}{2} \sum_{t=2}^n \ln \left\{ 1 + \frac{\Theta}{\nu} \left[ \Delta y_t - \frac{c y_{t-1}}{n} - \gamma' \Delta x_t + \frac{c \gamma' x_{t-1}}{n} \right]^2 \right\}$$

where the parameter  $\Theta$  measures the spread of the disturbance distribution and  $\nu$  is the degree of freedom that measures the tail thickness. Large values of  $\nu$  corresponds to thin tails in distribution. For given parameters  $\nu$  and  $\Theta$ , denoting  $\Theta/\nu$  as  $\theta$ , the MLE of  $(\gamma, c)$  is the solution of the following optimization problem

$$\begin{bmatrix} \bar{c} \\ \bar{\gamma} \end{bmatrix} = \arg \min \sum_{t=2}^n \ln \left\{ 1 + \theta \left[ \Delta y_t - \frac{c y_{t-1}}{n} - \gamma' \Delta x_t + \frac{c \gamma' x_{t-1}}{n} \right]^2 \right\}$$

The above optimizing problem can be solved as the following:

(i) for some appropriate choice  $\bar{c}$ , we calculate the estimator for  $\gamma$  under the hypothesis  $c = \bar{c}$ ,

$$\bar{\gamma} = \arg \min \sum_{t=2}^n \ln \left\{ 1 + \theta [\Delta_{\bar{c}} y_t - \gamma' \Delta_{\bar{c}} x_t]^2 \right\}. \quad (24)$$

(ii) Then, we construct the detrended data based on  $\bar{\gamma}$ ,  $\bar{y}_t^s = y_t - \bar{\gamma}' x_t$ , and re-estimating  $\bar{c}$  based on the M-estimator of the autoregressive coefficients of  $\bar{y}_t^s$ , i.e.,  $\hat{c} = n \hat{\phi}_o$ , where

$$\hat{\phi}_o = \arg \min \sum_{t=2}^n \ln \left\{ 1 + \theta [\Delta \bar{y}_t^s - \phi_o \bar{y}_{t-1}^s]^2 \right\} \quad (25)$$

If there is  $I(0)$  dependence in  $\mu_t$ , we replace 25 by



$$\begin{bmatrix} \hat{\phi}_o \\ \{\hat{\psi}_j\}_{j=1}^k \end{bmatrix} = \arg \min \sum_{t=2}^n \ln \left\{ 1 + \theta \left[ \Delta \bar{y}_t^s - \phi_o \bar{y}_{t-1}^s - \sum_{j=1}^k \psi_j \Delta y_{t-j}^s \right]^2 \right\}$$

where now we have  $\hat{c} = \frac{n\hat{\phi}_o}{1 - \sum_{j=1}^k \hat{\psi}_j}$ .

In practical analysis, the parameters  $\nu$  and  $\Theta$  are not known and has to be estimated. We consider a two-step partially adaptive estimator of  $(c, \gamma)$  in which the first step involves a preliminary estimation of the parameters  $\nu$  and  $\Theta$  (and thus  $\theta$ ). We then replace  $\theta$  in 25 by its estimator and perform a second step estimation for  $(c, \gamma)$ . In the presence of general disturbance distributions,  $\nu$  and  $\Theta$  lose their original meaning. However, in the cases where  $0 \leq \hat{\nu}$  and  $0 \leq \hat{\Theta}$ ,  $\hat{\nu}$  and  $\hat{\Theta}$  still can be interpreted as estimators of measures of the tailthickness and the spread of the disturbance distribution, and partially adaptive estimator 25 still have good sampling properties.

Potscher and Prucha (1986) discussed the estimation of the adaptation parameters  $\nu$  and  $\Theta$ . In particular, if we denote  $E(|u_t|^k)$  as  $\sigma_k$ , then for  $\nu > 2$ , we have

$$\frac{\sigma_2}{\sigma_1^2} = \frac{\pi}{\nu - 2} \frac{\Gamma[\nu/2]^2}{\Gamma[(\nu - 1)/2]^2} = \rho(\nu) \quad (26)$$

and

$$\Theta = \frac{1}{\pi} \frac{\nu \Gamma[(\nu - 1)/2]^2}{\sigma_1^2 \Gamma[\nu/2]^2} = q(\nu, \sigma_1).$$

Potscher and Prucha show that  $\rho(\cdot)$  is analytic and monotonically decreasing on  $(2, \infty)$  with  $\rho(2+) = \infty$  and  $\rho(\infty) = \pi/2$ . Thus, given estimator of  $\sigma_1$  and  $\sigma_2$ ,  $\nu$  can be estimated by inverting  $\rho(\nu)$  in 26 and thus an estimator of  $\theta$  can be obtained from

$$\hat{\theta} = \frac{q(\hat{\nu}, \hat{\sigma}_1)}{\hat{\nu}} = \frac{1}{\pi} \frac{\Gamma[(\hat{\nu} - 1)/2]^2}{\hat{\sigma}_1^2 \Gamma[\hat{\nu}/2]^2}.$$

For the estimation of  $\sigma_1$  and  $\sigma_2$ , we may use the sample moments

$$\hat{\sigma}_k = \frac{1}{n} \sum_t |\hat{u}_t|^k.$$

where  $\hat{u}_t$  is OLS estimator of  $u_t$ . Once we obtain the estimated (partially adaptive) density, the limiting distribution of  $\hat{c}$  can be approximated using simulation method. Thus, a partially adaptive unit root test can be conducted.

## 4 Monte Carlo Results.

We conducted a Monte Carlo experiments to examine the sampling performance of the M-detrended unit root test. In particular we focused on the size adjusted empirical power of the following unit root tests:

- (1) the Dickey-Fuller coefficient test based on OLS detrending, denoted by OLS;
- (2) the quasi-differencing DF test (ERS test), based on least square regression on the quasi-differenced data, denoted by QD;
- (3) the  $\bar{Z}_{QDM}$  test developed above and based on M-estimation plus quasi-differencing, denoted by QDM.

The model used for data generation was the following model:

$$\begin{aligned} y_t &= \gamma' x_t + y_t^s \\ y_t^s &= \alpha y_{t-1}^s + u_t, t = 1, \dots, n \end{aligned}$$

where the true value of  $\gamma$  is 0 and  $\{u_t\}$  is an iid sequence of t-distributions with three degrees of freedom. We standardized  $u_t$  so that it has unity variance. Two sample sizes were considered:  $n=100$ ,  $n=200$ . The number of iterations is 2,000 in each case, and the initial value of  $y^s$  is set at 0.

To provide a power comparison among the different tests, size corrected power is reported (for discussion on the use of size-corrected power, also see Stock, 1995; Cheung and Lai, 1997). The corresponding critical values were calculated from a direct simulation using 20,000 replications. Both the demeaned test and the detrended test are examined. In particular, Table I reports the empirical power of the demeaned tests, and Table II reports the power of the tests with a linear time trend.

Table I: Size-adjusted empirical power (demeaned case)

	OLS	QD	QDM
$\alpha=0.975$	0.0976	0.1002	0.1260
$\alpha=0.95$	0.1792	0.1822	0.2544
$\alpha=0.925$	0.2982	0.3076	0.4682
$\alpha=0.90$	0.4546	0.4688	0.7086
$\alpha=0.875$	0.6278	0.6416	0.8896
$\alpha=0.85$	0.7874	0.7962	0.9672
$\alpha=0.825$	0.8968	0.9016	0.9908
$\alpha=0.80$	0.9578	0.9606	0.9980

Table II: Size-adjusted empirical power (detrended case)

	OLS	QD	QDM
$\alpha=0.975$	0.0624	0.0638	0.0786
$\alpha=0.95$	0.1048	0.1068	0.1592
$\alpha=0.925$	0.1936	0.1950	0.3166
$\alpha=0.90$	0.3192	0.3222	0.5434
$\alpha=0.875$	0.4804	0.4822	0.7684
$\alpha=0.85$	0.6474	0.6522	0.9134
$\alpha=0.825$	0.7974	0.8024	0.9736
$\alpha=0.80$	0.8982	0.9006	0.9926

A general conclusion coming from these Monte Carlo experiments is that the testing procedure using distributional information has substantially improved power properties. In other words, our simulations indicate that the empirical power functions of test QDM is well above the power functions of least-square based tests (OLS and QD). This result together with the fact that the QDM unit root test based on the class of student-t distributions has a bounded influence function, suggests that we should test for unit root using the QDM approach whenever there is evidence of departures from Gaussianity in the data.

## 5 Re-visiting the PPP Hypothesis: Evidence From a Robust Approach

In the search for a more powerful unit root test, new studies on the PPP hypothesis have used the ERS test. In this section, we will re-visit two samples of real exchange rates by using our robust approach. We first consider the work of Taylor (2002). He uses a secular sample of US-dollar based real exchange rates<sup>4</sup> (1892 to 1996) that covers the set of countries Argentina, Australia, Belgium, Brazil, Canada, Denmark, Finland, France, Germany, Italy, Japan, Mexico, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and United Kingdom.

Figure 2 shows the graphics of real exchange rate for the countries listed above. It clearly suggest that almost all time series are characterized by the presence of outliers. The outliers appear to be additive: the series almost immediately returns to the normal level (trend) after the shock is over. The additive outliers in the majority of the European countries seem to be provoked by wars and oil shocks. Latin American countries like Argentina, Brazil and Mexico, experimented strong economic instability in the 80's and 90's and this may explain the AO's observed in those countries.

We now turn to unit root analysis. For comparison purposes, we considered the same lag choices and deterministic components as in Taylor (2002)<sup>5</sup> with the localizing parameter,  $\bar{c}$ , being set equal to  $-13.5$  for the specification with

<sup>4</sup>Taylor (2002) also considered real exchange rates relative to a world basket of currencies.

<sup>5</sup>To be precise, Taylor (2002) considered two specifications (intercept and linear trend) in all time series. To save space, just one specification was considered in this paper. The

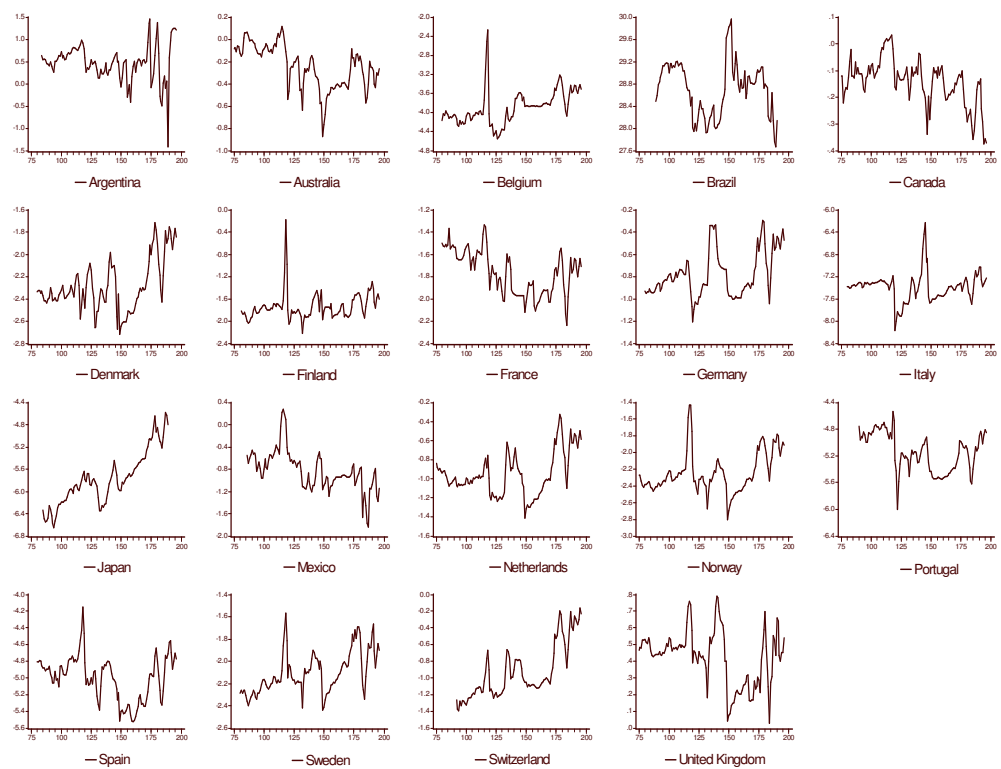


Figure 2: Real Exchange Rates (1892 to 1996)

a linear trend, and  $\bar{c} = -7$  for the specification with an intercept. The results are displayed in Table III. The test statistic of the quasi-difference M-detrended unit root test based on partial adaptive estimation is given by  $Z_{QDM}$ . Table III also shows estimates of the autoregressive coefficient  $\alpha$ . The first estimate,  $\hat{\alpha}_{ERS}$ , is coming from the ERS regression and, therefore, it is a OLS estimate of  $\alpha$ . The second one,  $\hat{\alpha}_{QDM}$ , corresponds to the partial adaptive estimator used in the QDM test. The high persistency in real exchange rates (and many other macroeconomic variables) suggests that the coefficient  $\alpha$  is near unity. However, if the series are contaminated by AO's, we have shown that the OLS estimator  $\hat{\alpha}_{ERS}$  will be biased towards zero whereas the M estimator based on Student-t distribution will have a bounded influence function. Finally, the last column includes the estimates the thickness parameter of the student-t distribution.

Unless for few countries, like Australia, France, Denmark, and Netherlands, the OLS estimator of the persistency parameter,  $\hat{\alpha}_{ERS}$ , seems to be smaller than the robust partial adaptive estimator,  $\hat{\alpha}_{QDM}$ , suggesting the presence of additive outliers in those series (a simple visual inspection of the graphics in Figure 2 reinforces this conclusion). Therefore, real exchange rates are more persistent than otherwise reported. Not surprisingly,  $Z_{ERS}$  is not significant for almost all countries, suggesting that PPP holds. When the robust test is applied, we find that the PPP still holds for many countries, but it fails in Belgium, Brazil, Canada, Japan, Norway, Portugal, Spain and Sweden. Finally, Table III shows estimates of the thickness parameter,  $\hat{\nu}$ . These estimates suggest that all time series analyzed display evidence of heavy-tailed distribution, which support the choice of the QDM rather than the ERS test to conduct unit root inference, since the former is more powerful than the latter under the presence of fat-tail distributions

---

specification with intercept was included for all countries but Canada, Japan, Mexico and Switzerland, in which we included a linear trend. Our choices reflect the trending behavior exhibited by the time series in Figure 1.

Table III: Unit Root Analysis (1892 - 1996)<sup>6</sup>

Countries	n	Test type	Lags	$Z_{ERS}$	$Z_{QDM}$	$\hat{\alpha}_{ERS}$	$\hat{\alpha}_{QDM}$	$\hat{\nu}$
Argentina	113	no trend	0	-37.00***	-13.0***	0.67	0.88	2.71
Australia	127	no trend	0	-12.00**	-11.90**	0.90	0.90	3.31
Belgium	117	no trend	0	-15.53***	-1.02	0.86	0.99	2.51
Brazil	108	no trend	1	-11.50**	-4.60	0.90	0.96	3.31
Canada	127	with trend	0	-8.72	-7.86	0.93	0.94	8.91
Denmark	117	no trend	0	-11.33**	-11.65**	0.90	0.90	3.91
Finland	116	no trend	0	-49.14***	-24.9***	0.70	0.80	2.51
France	117	no trend	1	-10.88**	-9.66**	0.93	0.94	8.91
Germany	117	no trend	1	-17.46***	-9.90**	0.92	0.95	3.31
Italy	117	no trend	0	-24.28***	-9.95***	0.83	0.92	2.51
Japan	105	with trend	0	-1.70	-5.30	0.97	0.95	4.51
Mexico	111	with trend	0	-23.94***	-13.19**	0.78	0.88	3.31
Netherlands	127	no trend	0	-9.04*	-7.23*	0.93	0.94	3.31
Norway	127	no trend	0	-5.24	-3.74	0.95	0.98	2.71
Portugal	107	no trend	0	-6.93*	-3.36	0.94	0.97	3.11
Spain	117	no trend	0	-8.44*	-5.12	0.92	0.96	4.51
Sweden	117	no trend	0	-10.13**	-4.32	0.91	0.96	3.11
Switzerland	105	with trend	1	-25.64***	-18.48***	0.85	0.89	7.11
UK	127	no trend	0	-14.40***	-8.14***	0.88	0.93	2.91

We now revisit the sample used by Cheung and Lai (1998). Their sample only includes post-Bretton wood era observations. In fact, the secular sample employed in Taylor (2002) has the advantage of pursuing a large time-span but, on the hand, it is criticized for comprising both flexible and fixed exchange rate regime periods, which, ultimately, may lead to rejection of the null of unit root due to low variability in the time series during the fixed regime periods. Cheung and Lai also employed the ERS test and find some support for PPP in the post-Bretton Woods era, but evidence was weak for US-dollar based real exchange rate. To construct the real exchange rate for the post-Bretton Woods era, we took the nominal exchange rate and the price level (Consumer Price Index) from the International Financial Statistics CD-ROM, which is made by the International Monetary Fund (IMF). The sample covers the same period as in Cheung and Lai (1998), from April 1973 to December 1994. Figure 2 shows the graphics of real exchange rate for the countries considered in Cheung and Lai, and there seem to be some evidence of additive outliers in the US-dollar based real exchange rates.

Table IV provides our unit root analysis. We considered the same lag and deterministic components choices as in Cheung and Lai (1998)<sup>7</sup>. First notice

<sup>6</sup>Finite-sample critical values were calculated based on 10,000 simulations of the null. The symbols (\*\*) and (\*) indicate rejection of the null hypothesis of unit root at 5% and 10% level of significance, respectively.

<sup>7</sup>Again, for comparison purposes with previous studies, we set the localizing parameter,  $\bar{\tau}$ , equal to  $-13.5$  for the specification with a linear trend, and  $\bar{\tau} = -7$  for the specification with an intercept.

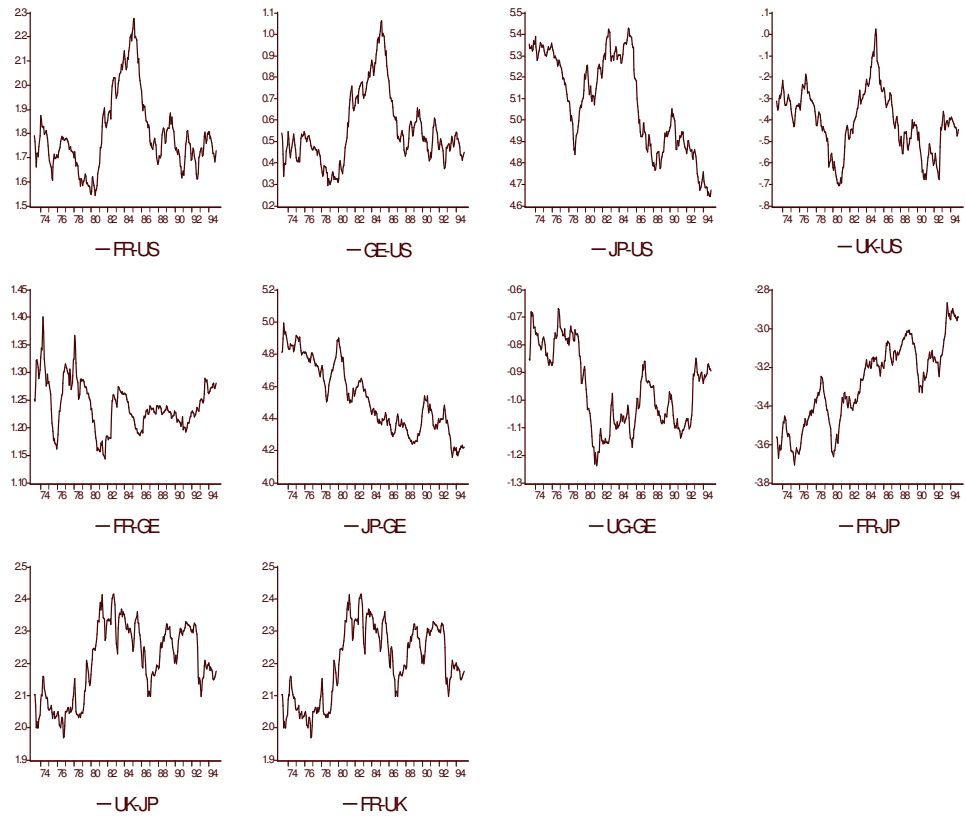


Figure 3: Real Exchange Rates (Post-Bretton Woods Era)

that the difference between  $\hat{\alpha}_{QDM}$  and  $\hat{\alpha}_{ERS}$  is more accentuated in the US-dollar based real exchange rates. This result matches with what is observed in Figure 3, that is, the real exchange rates relative to the US dollar seem to be contaminated by additive outliers. An interesting fact is that the difference between  $\hat{\alpha}_{QDM}$  and  $\hat{\alpha}_{ERS}$  is nil for JP/GE and FR/JP, which are the series that seems to have normal distribution according to the estimate of our thickness parameter<sup>8</sup>. This confirms the fact that the OLS and partial adaptive estimators will be equivalent under Gaussian innovations. This happens because the class of student-t distribution includes the normal distribution as a limit case. The statistic  $\bar{Z}_{ERS}$  lead us to the same conclusion reached by Cheung and Lai (1998): PPP holds for the majority of the countries in the sample (7 out of 10), but holds barely for the US-dollar RER. When the robust test  $\bar{Z}_{QDM}$  is employed and additive outliers are considered, we find that the PPP holds for only three (out of 10) countries, yielding little evidence in favor of PPP reversion. More specifically, the difference between our results and theirs resides in the US-dollar RER's, which are exactly the time series that seemed to be contaminated by outliers.

Table IV: Unit Root Analysis<sup>9</sup>

Countries	Test type	Lags	$Z_{ERS}$	$Z_{QDM}$	$\hat{\alpha}_{ERS}$	$\hat{\alpha}_{QDM}$	$\hat{\nu}$
FR/US	No trend	2	-6.44*	-4.56	0.982	0.991	10.11
GE/US	No trend	2	-6.70*	-5.95	0.981	0.988	19.91
JP/US	With trend	3	-9.37	-7.53	0.973	0.977	11.11
UK/US	No trend	2	-6.42*	-3.94	0.980	0.990	8.91
FR/GE	No trend	4	-15.64**	-7.67**	0.964	0.972	3.71
JP/GE	With trend	3	-24.7**	-25.6**	0.945	0.945	$\infty$
UK/GE	No trend	3	-5.64	-5.05	0.986	0.986	13.11
FR/JP	With trend	6	-33.60***	-37.24***	0.943	0.942	$\infty$
UK/JP	With trend	3	-12.66	-10.13	0.971	0.977	7.51
FR/UK	No trend	3	-6.26*	-5.09	0.983	0.986	6.11

## 6 Conclusion

This paper re-visited the PPP hypothesis using a robust approach and found less evidence in favor of PPP reversion than what was previously obtained using nonrobust unit root tests. We considered quasi-differencing coupled with M-estimation to construct a unit root test that is robust against two important features observed in real exchange rates: near-unity root and additive outliers. The PPP puzzle is more evident during the post-Bretton Woods era. It is important to mention that recent studies have found evidence in favor of PPP reversion using non-linear models. A partial list along this direction includes:

<sup>8</sup> Actually,  $\hat{\nu} = \infty$  means that the upper bound in our numerical optimatizatin was reached. We set the upper bound equal to 80.

<sup>9</sup> Finite-sample critical values were calculated based on 10,000 simulations of the null. The symbols (\*\*) and (\*) indicate rejection of the null hypothesis of unit root at 5% and 10% level of significance, respectively.



Chortareas and Kapetanios (2004), who find PPP reversion in the Yen real exchange rate by using a non-linear version of the augmented Dickey-Fuller test; Michael et al. (1997) also report evidence in favor of PPP hypothesis by considering nonlinear adjustment in real exchange rates; and Gouvea and Rodrigues (2004) do not reject the PPP hypothesis using the threshold cointegration adaptation of the unit root test procedure suggested by Caner & Hansen (2001). In this paper, we have not tested for nonlinearity, but we expect that the apparent nonlinearity in real exchange rates may be due to the existence of outliers. Indeed, Balke and Fomby (1994) and Van Dijk et al. (1999) show that, after controlling for outliers, much of the evidence of non-linearity in many economic time series is eliminated. In conclusion, our results not only suggest that further investigations on the PPP reversion are necessary to accumulate more evidence on the empirical relevance of the PPP hypothesis, but also emphasizes the importance of the use of robust methods in new studies on PPP in order to take into account the nonGaussian behavior observed in real exchange rate.

## 7 APPENDIX

### 7.1 Asymptotic properties of the M estimator

Xiao (2001) studied some asymptotic properties of the M estimator (17). He makes the following assumptions on  $u_t$  and the criterion function  $\varphi$ . These conditions are assumed for the convenience of asymptotic analysis. In practice, even if these conditions do not hold, as long as the data has similar distributional properties as the function  $\varphi$  described, Monte Carlo evidence indicates that the joint estimation still have good sampling properties.

ASSUMPTION A: The process  $\{u_t\}$  is strictly stationary and strong-mixing with mixing coefficients  $\alpha_m$  that satisfy  $\alpha_m = O(m^{-\lambda})$  for all  $\lambda > p\nu/(p-\nu)$  for some  $p > \nu > 2$ ;  $E(u_t) = 0$  and  $E|u_t|^p < \infty$ ;  $\omega_u^2 = \lim_{n \rightarrow \infty} E(\sum_{t=1}^n u_t)^2/n > 0$ . ASSUMPTION B:  $\varphi(\cdot)$  possesses derivatives  $\varphi'$  and  $\varphi''$ .  $[u, \varphi'(u)]$  has  $k$ -th moments for some  $k > 2$ ,  $E[\varphi'(u_t)] = 0$ , and  $\varphi''$  is Lipschitz continuous.

ASSUMPTION C:  $\tilde{c} = c + O_p(1)$ , and  $n^{-1/2}D_n(\tilde{\gamma} - \gamma) = o_p(n^{1/4})$ .

Assumptions A and B are standard conditions in asymptotic analysis of maximum likelihood estimators or M estimators. The mixing condition requires that the mixing coefficients should be of size  $p\nu/(p-\nu)$  (White 1984) so that the dependence of  $u_t$  and  $u_{t+k}$  is controlled. The moment conditions on  $u$  and  $\varphi'(u)$  and the mixing conditions are needed to establish the weak convergence results. We may also replace the moment condition on  $\varphi'(u)$  by boundness conditions of the derivatives of  $\varphi$ , because the latter and the moment condition on  $u$  imply the corresponding condition on  $\varphi'$ . Assumptions similar to C are also standard in the development of M estimator asymptotics. It is related to Assumption (b) in Theorem 5.1 of Phillips (1995) and the assumption on  $\tilde{\varepsilon}_t - \varepsilon_t$  in Theorem 1 of Lucas (1995).

Denote  $[\cdot]$  as the greatest lesser integer function. Then under Assumption A, as  $n$  goes to  $\infty$ ,  $n^{-1/2} \sum_1^{[nr]} u_t$  converges weakly to a Brownian motion  $B_1(r) = \omega_u W_1(r) = BM(\omega_u^2)$ , where  $\omega_u^2$  is the long run variance of  $u_t$ , denoted as  $\text{lrvar}(u_t)$ . Thus  $n^{-1/2} y_{[nr]}^s$  converges weakly to the corresponding Ornstein-Uhlenbeck process  $J_c(r) = \int_0^r e^{c(r-s)} dB_1(s)$ . The limiting distributions of  $\hat{c}$  and  $\hat{\gamma}$  will also be dependent on the weak limit of the partial sums of  $\varphi'(u_t)$ . Denoting  $\omega^2 = \text{lrvar}[\varphi'(u_t)]$ , and  $\delta = E[\varphi''(u_t)]$ , then  $n^{-1/2} \sum_1^{[nr]} \varphi'(u_t) \Rightarrow B_\varphi(r) = BM(\omega^2)$ .

For asymptotic analysis of the deterministic trend, we assume that there are standardizing matrices  $D_n$  and  $F_n = n^{-1} D_n$  such that  $D_n^{-1} x_{[nr]} \rightarrow X(r)$  and  $F_n^{-1} \Delta x_{[nr]} \rightarrow g(r)$ , as  $n \rightarrow \infty$ , uniformly in  $r \in [0, 1]$ . In the case of a linear trend,  $D = \text{diag}[1, n]$  and  $X(r) = (1, r)'$ . If  $x_t$  is a general  $p$ -th order polynomial trend,  $D = \text{diag}[1, n, \dots, n^p]$  and  $X(r) = (1, r, \dots, r^p)$ . Denoting the limit of  $n^{-1/2} D_n(\tilde{\gamma} - \gamma)$  by  $\xi_c$  and the limit of  $\hat{c}$  by  $\eta_c$ , the limiting distributions of the M estimators  $\hat{c}$  and  $\hat{\gamma}$  are given in the following theorem.

**THEOREM 1 (XIAO, 2001):** Given models (12), (13), and (15), for all  $c$  in a compact set, under Assumptions A, B, and C, the limiting distributions of nonlinear regression estimators  $\hat{\gamma}$  and  $\hat{c}$  are jointly determined by the following equations:

$$\begin{aligned}\xi_c &= \left[ \int X_\eta(r) X'_\eta(r) dr \right]^{-1} \int X_\eta(r) dS_c(r), \\ \eta_c &= \left[ \int \underline{J}_{c\xi}(r)^2 dr \right]^{-1} \left[ \int \underline{J}_{c\xi}(r) d\tilde{S}_c(r) + \lambda \right],\end{aligned}$$

where

$$\begin{aligned}X_\eta(r) &= g(r) - \eta_c X(r), \quad \underline{J}_{c\xi}(r) = J_c(r) - \xi'_c X(r), \\ \underline{J}_c(r) &= c J_c(r) - \xi'_c g(r), \quad \tilde{S}_c(r) = \int_0^r \underline{J}_c(s) ds + B_\varphi(r)/\delta, \\ S_c(r) &= (c - \eta_c) \int_0^r J_c(s) ds + B_\varphi(r)/\delta,\end{aligned}$$

and  $\lambda = \Delta_{u\varphi}/\delta$ ,  $\Delta_{u\varphi}$  = the one-sided long run covariance between  $u_t$  and  $\varphi'(u_t)$ .

## Proof

By definition, the estimators  $(\hat{c}, \hat{\gamma})$  solve the following equation system:

$$\begin{aligned}(i) : & \sum_{t=1}^n \varphi'(\Delta y_t - \tilde{c}(y_{t-1}/n) - \tilde{\gamma}' \Delta x_t + \tilde{c} \tilde{\gamma}'(x_{t-1}/n)) \Delta \tilde{c} x_t = 0, \\ (ii) : & \sum_{t=1}^n \varphi'(\Delta y_t - \tilde{c}(y_{t-1}/n) - \tilde{\gamma}' \Delta x_t + \tilde{c} \tilde{\gamma}'(x_{t-1}/n)) \tilde{y}_{t-1}^s = 0,\end{aligned}$$

where  $\tilde{y}_t^s = y_{t-1} - \gamma' x_{t-1}$ . For simplicity, we denote that

$$\hat{u}_t = \Delta y_t - \tilde{c}(y_{t-1}/n) - \tilde{\gamma}' \Delta x_t + \tilde{c} \tilde{\gamma}'(x_{t-1}/n);$$

then

$$\hat{u}_t = u_t - (\tilde{c} - c)(\tilde{y}_{t-1}^s/n) - (\tilde{\gamma} - \gamma)' \Delta_c x_t + (\tilde{c} - c)(\tilde{\gamma} - \gamma)'(x_{t-1}/n);$$

Under Assumption C, for all  $c$  in a compact set, the left-hand side of equation (i) is asymptotically equivalent to

$$(i') : \sum_{t=1}^n \varphi'(u_t) \Delta_{\tilde{c}x_t} - (\tilde{c} - c) \sum_{t=1}^n \varphi''(u_t) \left( \frac{\tilde{y}_{t-1}^s}{n} \right) \Delta_{\tilde{c}x_t} - (\tilde{\gamma} - \gamma) \sum_{t=1}^n \varphi''(u_t) \Delta_c x_t' \Delta_{\tilde{c}x_t} +$$

$$(\tilde{c} - c)(\tilde{\gamma} - \gamma) \sum_{t=1}^n \varphi''(u_t) \left( \frac{x_{t-1}'}{n} \right) \Delta_{\tilde{c}x_t}$$

Under the assumptions A, B and C, the following asymptotics hold:

$$F_n^{-1} \Delta_c x_{[nr]} \Rightarrow g(r) - \eta_c X(r) = X_\eta(r),$$

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} \varphi'(u_t) \Rightarrow B_\varphi(r) = \omega W_\varphi(r) = BM(\omega^2),$$

$$\frac{1}{\sqrt{n}} \tilde{y}_{[nr]}^s \Rightarrow J_c(r) - \xi_c' X(r) = \underline{J}_c \xi(r),$$

where  $J_c(r) = \int_0^r e^{c(r-s)} dB_1(s)$  and  $B_1(r)$  is the weak limit of  $n^{-1/2} \sum_{t=1}^n u_t$ .

As a result, it can be verified that

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \varphi'(u_t) \Delta_{\tilde{c}x_t} F_n^{-1} \Rightarrow \int_0^1 dB_\varphi X_\eta, \quad (27)$$

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \varphi''(u_t) (\tilde{c} - c) \left( \frac{\tilde{y}_{t-1}^s}{n} \right) \Delta_{\tilde{c}x_t} F_n^{-1} \Rightarrow \delta(\eta_c - c) \int_0^1 J_c X_\eta, \quad (28)$$

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \varphi''(u_t) (\tilde{\gamma} - \gamma)' \Delta_c x_t' \Delta_{\tilde{c}x_t} F_n^{-1} \Rightarrow \delta \xi_c' \int_0^1 X_c X_\eta. \quad (29)$$

$$\frac{1}{\sqrt{n}} \sum_{t=1}^n \varphi''(u_t) (\tilde{c} - c)(\tilde{\gamma} - \gamma)' \left( \frac{x_{t-1}'}{n} \right) \Delta_{\tilde{c}x_t} F_n^{-1} \Rightarrow \delta(\eta_c - c) \xi_c' \int_0^1 X X_\eta \quad (30)$$

Denote

$$S_c(r) = (c - \eta_c) \int_0^r J_c(s) ds + B_\varphi(r)/\delta, \text{ and } \tilde{S}_c(r) = \int_0^r \underline{J}_c(s) ds + B_\varphi(r)/\delta.$$

Then, by the results of 27-30 and (i'), we have

$$\xi_c = \left[ \int X_\eta(r) X_\eta'(r) dr \right]^{-1} \int X_\eta(r) dS_c(r).$$

Similarly, under the given assumptions, the left-hand side of equation (ii) is asymptotically equivalent to

$$(ii') : \frac{1}{\sqrt{n}} \sum_{t=1}^n \varphi'(u_t) \tilde{y}_{t-1}^s - (\tilde{c} - c) \sum_{t=1}^n \varphi''(u_t) \left( \frac{\tilde{y}_{t-1}^s}{n} \right) \tilde{y}_{t-1}^s - (\tilde{\gamma} - \gamma)' \sum_{t=1}^n \varphi''(u_t) \Delta_c x_t \tilde{y}_{t-1}^s +$$

$$(\tilde{c} - c)(\tilde{\gamma} - \gamma)' \sum_{t=1}^n \varphi''(u_t) \left( \frac{x_{t-1}'}{n} \right),$$

and by calculations of weak limits of the sample covariances  $(1/n) \sum_{t=1}^n \varphi'(u_t) \tilde{y}_{t-1}^s$ ,

$$(1/n^2) \sum_{t=1}^n \varphi''(u_t) \tilde{y}_{t-1}^s \tilde{y}_{t-1}^s, (1/n)(\tilde{\gamma} - \gamma)' \sum_{t=1}^n \varphi''(u_t) \Delta_c x_t \tilde{y}_{t-1}^s, \text{ and } (1/n)(\tilde{\gamma} - \gamma)' \sum_{t=1}^n \varphi''(u_t) \left( \frac{x_{t-1}'}{n} \right) \tilde{y}_{t-1}^s,$$

following a similar argument as the previous proof, we get

$$\eta_c = \left[ \int \underline{J}_c \xi(r)^2 dr \right]^{-1} \left[ \int \underline{J}_c \xi(r) d\tilde{S}_c(r) + \lambda \right].$$

## Proof of Proposition 1

We only need to employ the arguments used above plus equations 20 and 21.

## References

- [1] Abauf, N., and P. Jorion (1990). Purchasing power parity in the long run. *Journal of Finance* 45, 157-174
- [2] Balke, N.S., and T.B. Fomby (1994). Large shocks, small shocks, and economic fluctuations: outliers in macroeconomic time series. *Journal of Applied Econometrics* 9, 181-200.
- [3] Beelders, O. (1998) Adaptive estimation and unit root tests. *Working Paper*. Emory University.
- [4] Bickel, P. (1982). On adaptive estimation. *Annals of Statistics*, 10, 647-671.
- [5] Caner, M. and B. E. Hansen (2001). Threshold autoregression with a unit root. *Econometrica* 69, 1555-1596.
- [6] Cheung, Y.W. and K.S. Lai (1997). Bandwidth selection, prewhitening, and the power of the Phillips-Perron test. *Econometric Theory* 13, 679-691
- [7] Cheung, Y.W., and K. S.Lai.(1998) Parity reversion in real exchange rates during the post-Bretton Woods period. *Journal of International Money and Finance* 17, 597-614.
- [8] Chortareas, G. and G. Kapetanios (2004) The Yen Real Exchange Rate May Be Stationary After All: Evidence from Non-linear Unit-Root Tests. *Oxford Bulletin of Economics and Statistics*, v66, 113-31
- [9] Cox, D. and I. Llatas (1991) Maximum likelihood type estimation for nearly nonstationary autoregressive time series. *Annals of Statistics* 19, 1109-1128.
- [10] Diebold, F.X., Husted, S., and Rush, M. (1991) Real exchange rates under the gold standard. *Journal of Political Economy* 99, 1252-1271.
- [11] Elliot, G., T.J. Rothenberg, and J.H. Stock (1996) Efficient tests for an autoregressive unit root. *Econometrica* 64, 813-836.
- [12] Frankel, J.A. (1986) International capital mobility and crowding-out in the US economy: imperfect integration of financial markets or goods markets? In: Hafer, R.W. (ed.), *How Open is the U.S. Economy? Lexington Books*, Lexington, pp. 33-67.
- [13] Frankel, J.A., and A.K. Rose, (1996) A panel project on purchasing power parity: mean reversion within and between Countries. *Journal of International Economics* 40, 209-224.
- [14] Froot, K.A., and K. Rogoff (1995) Perspectives on PPP and long-run real exchange rates, In: Grossman, G., Rogoff, K. (eds.), *Handbook of International Economics*, vol. 3, North-Holland, New York, pp. 1647-1688.

- [15] Gouveia, P. M. D. C. B., and P. M. M. Rodrigues. ( 2004) Threshold Cointegration and the PPP Hypothesis. *Journal of Applied Statistics*, 31, 115-27.
- [16] Grenander, U., and M. Rosenblat (1957) *Statistical Analysis of Stationary Time Series*. New York: John Wiley.
- [17] Hall, D. and B. Joiner. (1982) Representations of the space of distributions useful in robust estimation of location. *Biometrika*, 69, 55-59.
- [18] Hansen, B. and S. Lee. (1994) Asymptotic Theory for Garch(1,1) quasi-maximum likelihood estimator. *Econometric Theory*, 10, 29-52.
- [19] Hoek, H., A. Lucas and H. K. van Dijk (1996) Classical and Bayesian aspects of robust unit root inference. *Journal of Econometrics*, 69, 27-59.
- [20] Huber, P.J. (1964) Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35, 73-101.
- [21] Huber, P.J. (1973) Robust regression: asymptotics, conjectures, and monte carlo. *Annals of Statistics*, 1, 799-821.
- [22] Im, K.S., M.H. Pesaran and Y. Shin (1996). Testing for unit roots in heterogeneous panels. Discussion paper, University of Cambridge.
- [23] Koenker, R., and Z. Xiao (2003). Testing stationarity using M-estimation. Working paper, University of Illinois at Urbana-Champaign.
- [24] Knight, K. (1991) Limiting Theory for M estimators in an integrated finite variance process. *Econometric Theory*, 7, 200-212.
- [25] Levin, A. and C.F. Lin (1992). Unit root tests in Panel Data: Asymptotic and finite sample properties. University of California at San Diego, Discussion Paper # 92-93.
- [26] Levin, A. and C.F. Lin (1993). Unit root tests in Panel Data: new results. University of California at San Diego, Discussion Paper # 93-56.
- [27] Lothian, J.R., and Taylor, M.P.(1996) Real exchange rate behavior: the recent float from the perspective of the past two centuries. *Journal of Political Economy* 104, 488-509.
- [28] Lothian, J. R. (1990). A century plus of yen exchange rate behavior. *Japan and the World Economy* 2, 47-70.
- [29] Lucas, A. (1994). An outlier robust unit root test with an application to the extended Nelson-Plosser data. *Journal of Econometrics* 66, 153-174.
- [30] Lucas, A. (1995). Unit root tests based on M estimators. *Econometric Theory* 11, 331-346.

- [31] MacDonald, R. (1985) Do deviations of the real effective exchange rate follow a random walk?. *Economic Notes*, 14(1), 63-9.
- [32] MacDonald, R. (1995). Random walks, real exchange rates and panel unit root tests. *Economic Letters*.
- [33] McCloskey, D.N., and J.R. Zecher (1984) The success of purchasing power parity: historical evidence and its implications for macroeconomics. In M. D. Bordo and A. J. Schwartz (Eds), *A retrospective of the classical gold standard, 1821-1931* (Chicago: University of Chicago Press)
- [34] Maddala, G. S. and S. Wu (1998). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*. (*forthcoming*).
- [35] Maddala, G.S., S. Wu and P.C. Liu (2000). Do panel data rescue the purchasing power parity (PPP) theory? In J. Krishnakumar and E. Ronchetti (Eds), *Panel Data Econometrics: Future Directions*, 35-51 (North-Holland).
- [36] Michael, Panos, A. R. Nobay, and D. A. Peel (1997). Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: An Empirical Investigation. *Journal of Political Economy* 105, 862-79.
- [37] Oh, K. Y. (1996). Purchasing Power Parity and the Unit Root Tests Using Panel Data. *Journal of International Money and Finance*, 15, 405-418.
- [38] Papell, D.H. (1997) Searching for stationarity: purchasing power parity under the current float. *Journal of International Economics* 49, 269-288.
- [39] Phillips, P.C.B. (1995). Robust nonstationary regression. *Econometric Theory* 12, 912-951.
- [40] Phillips, P.C.B. and Z. Xiao (1998). A primer on unit root testing. *Journal of Economic Survey* 12, 423-469.
- [41] Potscher, B., I. Prucha. (1986) A class of partially adaptive one-step M estimators for the nonlinear regression model with dependent observations. *Journal of Econometrics* 32, 219-251
- [42] Rogoff, K. (1996) The purchasing power parity puzzle. *Journal of Economic Literature* 34, 647-668.
- [43] Rothenberg, T., and J. Stock (1997) Inference in a nearly integrated autoregressive model with nonnormal innovations. *Journal of Econometrics* 80, 269-286.
- [44] Seo, B. Essays on Time Series Econometrics. *PhD Dissertation*, University of Rochester.

- [45] Schmidt, P. and P.C.B. Phillips (1992) Testing for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, 54, 257-288.
- [46] Stock, J.H. (1995) Unit roots, structural breaks, and trends. In R.F. Engle and Mcfadden (eds.), *Handbook of Econometrics*, vol.4, pp. 2739-2841. Amsterdam: North-Holland.
- [47] Taylor, A.M. (2002). A century of purchasing-power parity. *Review of Economics and Statistics* 84, 139-150.
- [48] Taylor, M. P. and L. Sarno (1998). The behavior of real exchange rates during the post-Bretton Woods period. *Journal of International Economics* 46, 281-312.
- [49] Van Dijk, D., P. H. Franses, and A. Lucas (1999) Testing for smooth transition nonlinearity in the presence of outliers. *Journal of Business & Economic Statistics*, 17, 217-234.
- [50] White, H. and I. Domowitz (1984). Nonlinear regression with dependent observations. *Econometrica*, 52, 143-161.
- [51] Xiao, Zhijie, (2001) Likelihood Based Inference for Trending Time Series with a Root Near Unity, *Econometric Theory*.

# ENSAIOS ECONÔMICOS DA EPGE

502. NONPARAMETRIC SPECIFICATION TESTS FOR CONDITIONAL DURATION MODELS - Marcelo Fernandes; Joachim Grammig – Setembro de 2003 – 42 págs.
503. A NOTE ON CHAMBERS'S "LONG MEMORY AND AGGREGATION IN MACROECONOMIC TIME SERIES" – Leonardo Rocha Souza – Setembro de 2003 – 11 págs.
504. ON CHOICE OF TECHNIQUE IN THE ROBINSON-SOLOW-SRINIVASAN MODEL - M. Ali Khan – Setembro de 2003 – 34 págs.
505. ENDOGENOUS TIME-DEPENDENT RULES AND THE COSTS OF DISINFLATION WITH IMPERFECT CREDIBILITY - Marco Bonomo; Carlos Viana de Carvalho – Outubro de 2003 – 27 págs.
506. CAPITAIS INTERNACIONAIS: COMPLEMENTARES OU SUBSTITUTOS? - Carlos Hamilton V. Araújo; Renato G. Flôres Jr. – Outubro de 2003 – 24 págs.
507. TESTING PRODUCTION FUNCTIONS USED IN EMPIRICAL GROWTH STUDIES - Pedro Cavalcanti Ferreira; João Victor Issler; Samuel de Abreu Pessoa – Outubro de 2003 – 8 págs.
508. SHOULD EDUCATIONAL POLICIES BE REGRESSIVE ? Daniel Gottlieb; Humberto Moreira – Outubro de 2003 – 25 págs.
509. TRADE AND CO-OPERATION IN THE EU-MERCOSUL FREE TRADE AGREEMENT - Renato G. Flôres Jr. – Outubro de 2003 – 33 págs.
510. OUTPUT CONVERGENCE IN MERCOSUR: MULTIVARIATE TIME SERIES EVIDENCE - Mariam Camarero; Renato G. Flôres Jr; Cecílio Tamarit – Outubro de 2003 – 36 págs.
511. ENDOGENOUS COLLATERAL - Aloísio Araújo; José Fajardo Barbachan; Mario R. Páscoa – Novembro de 2003 – 37 págs.
512. NON-MONOTONE INSURANCE CONTRACTS AND THEIR EMPIRICAL CONSEQUENCES - Aloísio Araujo; Humberto Moreira – Novembro de 2003 – 31 págs.
513. EQUILIBRIA IN SECURITY MARKETS WITH A CONTINUUM OF AGENTS - A. Araujo; V. F. Martins da Rocha; P. K. Monteiro – Novembro de 2003 – 17 págs.
514. SPECULATIVE ATTACKS ON DEBTS AND OPTIMUM CURRENCY AREA: A WELFARE ANALYSIS - Aloisio Araujo; Márcia Leon – Novembro de 2003 – 50 págs.
515. O MÉTODO GENERALIZADO DOS MOMENTOS(MGM): CONCEITOS BÁSICOS - Renato G. Flôres Jr – Novembro de 2003 – 27 págs.
516. VARIÁVEIS INSTRUMENTAIS E O MGM: USO DE MOMENTOS CONDICIONAIS - Renato G. Flôres Jr – Novembro de 2003 – 27 págs.
517. O VALOR DA MOEDA E A TEORIA DOS PREÇOS DOS ATIVOS - Fernando de Holanda Barbosa – Dezembro de 2003 – 17 págs.
518. EMPRESÁRIOS NÁVICOS, GARANTIAS E ACESSO À CRÉDITO - Marcelo Côrtes Néri; Fabiano da Silva Giovanini - Dezembro de 2003 – 23 págs.



519. DESENHO DE UM SISTEMA DE METAS SOCIAIS - Marcelo Côrtes Néri; Marcelo Xerez - Dezembro de 2003 – 24 págs.
520. A NEW INCIDENCE ANALYSIS OF BRAZILIAN SOCIAL POLICIES USING MULTIPLE DATA SOURCES - Marcelo Côrtes Néri - Dezembro de 2003 – 55 págs.
521. AN INTRA-HOUSEHOLD APPROACH TO THE WELFARE COSTS OF INFLATION - Rubens Penha Cysne – Janeiro de 2004 – 16 págs.
522. CENTRAL LIMIT THEOREM FOR ASYMMETRIC KERNEL FUNCTIONALS - Marcelo Fernandes; Paulo Klinger Monteiro – Fevereiro de 2004 – 23 págs.
523. THE TRADE-OFF BETWEEN INCENTIVES AND ENDOGENOUS RISK - Aloísio Araujo; Humberto Moreira; Marcos H. Tsuchida – Fevereiro de 2004 – 21 págs.
524. DO DIVIDENDS SIGNAL MORE EARNINGS ? - Aloísio Araujo; Humberto Moreira; Marcos H. Tsuchida – Fevereiro de 2004 – 26 págs.
525. Biased managers, organizational design, and incentive provision - Cristiano M. Costa; Daniel Ferreira; Humberto Moreira – Fevereiro de 2004 – 11 págs.
526. Land taxes in a Latin American context - Juliano J. Assunção; Humberto Moreira – Fevereiro de 2004 - 19 págs.
527. Indicadores coincidentes de atividade econômica e uma cronologia de recessões para o Brasil - Angelo J. Mont'alverne Duarte; João Victor Issler; Andrei Spacov - Fevereiro de 2004 – 41 págs.
528. TESTING UNIT ROOT BASED ON PARTIALLY ADAPTIVE ESTIMATION - Zhijie Xiao; Luiz Renato Lima – Março de 2004 – 27 págs.
529. DO SHOCKS PERMANENTLY CHANGE OUTPUT? LOCAL PERSISTENCY IN ECONOMIC TIME SERIES - Luiz Renato Lima; Zhijie Xiao – Março de 2004 – 21 págs.
530. A NEW PERSPECTIVE ON THE PPP HYPOTHESIS - Soyoung Kim; Luiz Renato Lima – Março de 2004 – 36 págs.
531. TRADE LIBERALIZATION AND INDUSTRIAL CONCENTRATION: EVIDENCE FROM BRAZIL - Pedro Cavalcanti Ferreira; Giovanni Facchini – Março de 2004 - 25 págs.
532. REGIONAL OR EDUCATIONAL DISPARITIES? A COUNTERFACTUAL EXERCISE - Angelo José Mont'Alverne; Pedro Cavalcanti Ferreira; Márcio Antônio Salvato – Março de 2004 – 25 págs.
533. INFLAÇÃO: INÉRCIA E DÉFICIT PÚBLICO – Fernando de Holanda Barbosa – Março de 2004 – 16 págs.
534. A INÉRCIA DA TAXA DE JUROS NA POLÍTICA MONETÁRIA – Fernando de Holanda Barbosa – Março de 2004 – 13 págs.
535. DEBT COMPOSITION AND EXCHANGE RATE BALANCE SHEET EFFECTS IN BRAZIL: A FIRM LEVEL ANALYSIS - Marco Bonomo; Betina Martins ; Rodrigo Pinto – Março de 2004 – 39 págs.

536. THE SET OF EQUILIBRIA OF FIRST-PRICE AUCTIONS - Paulo Klinger Monteiro – Março de 2004 – 6 págs.
537. OPTIMAL AUCTIONS WITH MULTIDIMENSIONAL TYPES AND THE DESIRABILITY OF EXCLUSION - Paulo Klinger Monteiro ; Benar Fux Svaiter; Frank H. Page Jr – Março de 2004 – 8 págs.
538. INCOME INEQUALITY IN A JOB-SEARCH MODEL WITH HETEROGENEOUS TIME PREFERENCES – Rubens Penha Cysne – Março de 2004 – 35 págs.
539. IMPOSTO INFLACIONÁRIO E TRANSFERÊNCIAS INFLACIONÁRIAS NO BRASIL: 1947-2003 - Rubens Penha Cysne; Paulo C. Coimbra-Lisboa – Março de 2004 – 7 págs.
540. ON THE STATISTICAL ESTIMATION OF DIFFUSION PROCESSES – A survey - Rubens Penha Cysne – Abril de 2004 – 26 págs.
541. ROBUSTNESS OF STATIONARY TESTS UNDER LONG-MEMORY ALTERNATIVES - Luiz Renato Lima; Zhijie Xiao – Abril de 2004 – 23 págs.
542. MONETARY UNION AND PRODUCTIVITY DIFFERENCES IN MERCOSUR COUNTRIES – Mariam Camarero; Renato G. Flôres, Jr.; Cecilio R. Tamarit – Abril de 2004 – 21 págs.
543. TWO ADDITIONS TO LUCAS ´S “INFLATION AND WELFARE” - Rubens Penha Cysne – Abril de 2004 – 9 págs.
544. THE IMPLICATIONS OF EMBODIMENT AND PUTTY-CLAY TO ECONOMIC DEVELOPMENT - Samuel de Abreu Pessoa; Rafael Rob – Abril de 2004 – 70 págs.
545. MONEY WITH BANK NETWORKS - Ricardo Cavalcanti; Henrique Dezemone Forno – no prelo.
546. CYCLICAL INTEREST PAYMENTS ON INSIDE MONEY - Ricardo Cavalcanti; Henrique Dezemone Forno – no prelo.
547. DOIS EXPERIMENTOS DE POLÍTICA MONETÁRIA NO MODELO NOVO-KEYNESIANO – Fernando de Holanda Barbosa – Abril de 2004 – 9 págs.
548. THE EVOLUTION OF INTERNATIONAL OUTPUT DIFFERENCES (1960-2000): FROM FACTORS TO PRODUCTIVITY - Pedro Cavalcanti Ferreira; Samuel de Abreu Pessoa; Fernando A. Veloso – Junho de 2004 – 31 págs.
549. PRINCIPAIS CARACTERÍSTICAS DO CONSUMO DE DURÁVEIS NO BRASIL E TESTES DE SEPARABILIDADE ENTRE DURÁVEIS E NÃO-DURÁVEIS - Fábio Augusto Reis Gomes; João Victor Issler – Márcio Antônio Salvato – Junho de 2004 - 28 págs.
550. AVALIANDO PESQUISADORES E DEPARTAMENTOS DE ECONOMIA NO BRASIL A PARTIR DE CITAÇÕES INTERNACIONAIS - João Victor Issler; Rachel Couto Ferreira - Junho de 2004 - 34 págs.
551. COMMON AGENCY WITH INFORMED PRINCIPALS - David Martimort; Humberto Moreira – Junho de 2004 – 37 págs.
552. PURCHASING POWER PARITY AND THE UNIT ROOT TESTS: A ROBUST ANALYSIS - Zhijie Xiao; Luiz Renato Lima – Julho de 2004 – 30 págs.